LINKS BETWEEN FREIGHT TRIP GENERATION RATES, ACCESSIBILITY AND SOCIO-DEMOGRAPHIC VARIABLES IN URBAN ZONES

Jesus GONZALEZ-FELIU¹, Laura PALACIOS-ARGÜELLO², Carlos SUAREZ-NUÑEZ³

¹ Department of Supply Chain, Purchasing and Project Management, Center of Research in Managerial Intelligence and Innovation - CERIIM, Excelia Group La Rochelle Business School, La Rochelle, France
² Department of Environment and Organization Engineering, Institut Henri Fayol, Ecole des Mines de Saint-Etienne, France
³ Institute of Innovation in Productivity and Logistics CATENA-USFQ, Department of Industrial Engineering, Universidad San Francisco de Quito, Ecuador

Abstract:
This paper proposes an assessment of the links between freight trip generation (FTG) rates and accessibility. First, the paper overviews the background, sets the context and motivates the research. Second, it presents the proposed methodology, which combines an FTG model, two accessibility indicators and a linear regression analysis to assess the relationships between freight trip demand and a set of socio-demographic variables including accessibility. The FTG modelling framework, adapted from previous works, allows estimating the number of freight trips with a small amount of standard data, even when no surveyed data is available. The two gravity accessibility indexes, one potential and one exponential, are defined in the continuity of recent freight accessibility works. To those indicators, a set of socio-demographic variables, including population, area or a zone (or density), are introduced. The relationships between FTG and all those variables are assessed via standard linear regression methods completed by the verification of the corresponding linear relationship hypotheses. Third, the framework is applied to the urban area of Lyon (France), where no urban goods survey data is available. Results show that potential accessibility seems to have a better correlation to FTG and could be a good decision support indicator when combined with the population as an explanatory variable. The population can be added to accessibility as an explanatory variable, the resulting models with two variables have a slightly lower accuracy but remains close to that of models with only accessibility as an explanatory variable. This work remains exploratory and finishes by proposing practical implications and further development lines.

Keywords: urban goods transport modelling, city logistics, land use, FTG, accessibility, decision support

To cite this article:

Contact:
1) gonzalezfeliu@excelia-group.com [https://orcid.org/0000-0002-0056-7627], 2) [https://orcid.org/0000-0002-9512-1344], 3) [https://orcid.org/0000-0001-8375-1843]

Article is available in open access and licensed under a Creative Commons Attribution 4.0 International (CC BY 4.0)
1. Introduction

Freight demand modelling is one of the main issues of urban logistics (Comi et al., 2012). Indeed, the estimation of freight trips related to each establishment of an urban zone is crucial for making diagnoses or for supporting decision-making methods that need this demand to optimize transport supply (Comi et al., 2018; Moufad and Jawab, 2019). This is done, in general, by using integrated freight models that allow to estimate freight trip and/or commodity origin-destination (O/D) matrices, as well as travelled distances and travel/stop times, among others. In this context, we observe in the literature a plethora of approaches that lead to different authors to produce various attempts of classification and typology frameworks (Comi et al., 2012; Anand et al., 2015; Gonzalez-Feliu, 2019), each having its own specificities but all works coinciding on the non-existence of a unified or standard framework.

Nevertheless, most urban freight transport demand models approaches have one common point: they need a generation phase (Gonzalez-Feliu, 2019). That generation phase is essential in the modelling process since it defines the flow quantities and characteristics, compared to the remaining phases that are mainly related to their segmentation and spatial distribution (Ortuzar and Willumsen, 2001). Therefore, this paper’s main focus is on this generation phase.

Furthermore, urban transport planning and engineering issues remain mainly related to personal mobility, relying, among others, on analyses of accessibility, for which various indicators and assessment frameworks are nowadays available (van Wee, 2016). Those frameworks allow to relate the facility to access to destinations from an origin to personal trips, estimated either via surveys or specific passenger transport models (Ortuzar and Willumsen, 2001). Although the joint use of accessibility indicators and transport models is common in personal mobility issues, this is not the case in urban goods transport, a field in which accessibility is less used (van Wee, 2016) and does seldom rely on freight transport flow estimation (Gonzalez-Feliu, 2019).

On another hand, the need of integrating freight and passenger trips is shown in various key fields of engineering, such as traffic modelling and simulation, land-use analysis and engineering, industrial and commercial zone development, infrastructure planning, building construction or logistics strategic planning, among others. Indeed, relating Freight Trip Generation (FTG) to accessibility appears, analogously than what it is seen in personal mobility, can support the integration of passenger and freight flows in global analyses of urban mobility and define systemic assessment methods that consider those flows and their interactions. Nevertheless, to the best of our knowledge, the analysis of the links between freight accessibility and freight trip generation remain underexplored. Thus, this paper also aims to examine this relationship, and more precisely the potential of representing FTG with a function of accessibility. First, the proposed methodology is presented, which includes a generation phase, an accessibility estimation stage and a linear regression analysis. Second, and to illustrate the methodology, an example of an application in Lyon (France) is proposed. The assessment results are presented and discussed. Finally, and as a conclusion, the paper addresses the main research and practice implications of the proposed framework.

2. Literature review

Freight demand modelling is one of the main issues of urban logistics, as shown by a plethora of works dealing to estimate urban goods flows (Comi et al., 2012; Anand et al., 2015; Gonzalez-Feliu, 2019). Despite that amount and variety of approaches, most models start by a generation phase (Gonzalez-Feliu, 2019), i.e. a phase where flows are estimated, either at the origin, at the destination, or at both, as on classical 4-step models of personal mobility (Ortuzar and Willumsen, 2001). Only origin-destination synthesis models define directly O/D matrices without passing through a classical generation phase (Sánchez-Díaz et al., 2015), but those approaches are only used in cases where almost only traffic counts are available as data for modelling, and the O/D synthesis methodology has several common aspects with generation.

For those reasons, this paper focuses on generation models. Although the approaches and units can be different, all of them relate to one of the two following elementary variables: the vehicle stop (which is the destination of a trip and the origin of another one) or a commodity quantity being delivered or picked up (expressed in weight, volume, surface or cost). Indeed, although we observe five modelling
units related to urban goods transport\textsuperscript{1}, at an elementary stage, all models are related either to the vehicle stop (trips, deliveries, routes then decomposed in trips and or stops, etc.) or to the quantity of the commodity to deliver (González-Feliu and Sánchez-Díaz, 2019). Two main categories of models are then related to those units: Freight Trip Generation (FTG) in the first case, Freight Generation (FG) in the second (Holguín-Veras et al., 2011). Concerning the methodological frameworks, categorical freight trip generation (FTG) (Bastida and Holguín-Veras, 2009) remains nowadays as the dominating type of generation models, and has been applied to different cities and by different research teams (Bastida and Holguín-Veras., 2009; Sánchez-Díaz et al., 2016; González-Calderérón et al., 2018; Holguín-Veras et al., 2018; González-Feliu and Sánchez-Díaz, 2019). That generation modelling type is based on the definition and use of categories; for each of them, a generation logic is identified and modelled. Categorical FTG (Holguín-Veras et al., 2011) can be mainly done via constant generation or functional form approaches (Gonzalez-Feliu and Sanchez-Diaz, 2019). Constant generation assigns a constant number of freight trips to each category (Ahrens et al., 1977; Le Nir and Routhier, 1995, González-Calderérón et al., 2018). Functional form modelling proposes different generation patterns, constant, linear or non-linear, as a function of the category (Holguín-Veras et al., 2011, 2013; Sánchez-Díaz et al., 2016; Caspersen, 2018; González-Feliu and Sánchez-Díaz, 2019). In such models, the identification of the most suitable functional form (i.e. the formalization of the relationship between the explained variable and the potential explanatory ones) is more important than the data granularity choice when defining categorical FTG models, but depends strongly on the data quality and availability. For those reasons, functional form models have nowadays become the dominating approach for FTG (Holguin-Veras et al., 2018).

As said in the introduction, in personal mobility engineering and planning, trip generation is often associated with other indicators like accessibility (Geurs and van Wee, 2004). Indeed, those indicators are usually applied in transport and land use planning and many authors have studied the relationships between personal trip generation and accessibility (Hanson and Schwab, 1987; Boarnet and Crane, 2001; Geurs and van Wee, 2004; van Wee, 2016). Accessibility is then a useful tool not only in land-use planning and urban engineering, for example in real estate and urban construction planning (to define priority zones for residential or commercial building construction), in retailing development and in civil engineering (for infrastructure and parking needs), but also in transportation engineering (mainly for public/private passenger transport planning and management). However, and according to van Wee (2016), freight accessibility constitutes a real need for practitioners and starts to be considered in research. For example, the development of decision support methods (Crainic et al., 2010; Galkin, 2017, Comi et al, 2018) can rely on such indicator. We find also some works of socio-economic nature using freight accessibility, in different forms, as a measure (Giuliano et al., 2016).

If we consider both the form and use of the indicators, we can divide such works into two main categories. The first includes works dealing with freight accessibility that assess the suitability of potential terminal locations or the development of infrastructures. In those works, accessibility takes mainly the form of traffic congestion indicators (Yachiyo Engineering Co. LTD and Pacific Consultants International, 2005; Chiabaut et al., 2016), distance accessibility (Thomas et al., 2003; Rodrigue, 2004; Bowen et al., 2008) or time/cost indicators (Kotavaara et al., 2017; Ibarra-Rojas et al., 2018). The second group of works defines exponential gravity-based indicators in order to analyze the suitability of location and distribution systems, mainly related to retailing (Crainic et al., 2010; Gonzalez-Feliu, 2019) or for the analysis of socio-economic characteristics of production, logistics or retailing zones (Helling, 1998; Giuliano et al., 2010, 2016; van den Heuvel et al., 2014; Gonzalez-Feliu and Peris-Pla, 2017).

In all these works, accessibility indicators, related to zones, are mainly used to analyze the suitability of defining infrastructures, land-use policies or logistics systems, but, to the best of our knowledge, the relationship between freight transport demand and accessibility remains to be explored. However, in

\textsuperscript{1} According to Gonzalez-Feliu (2019): the vehicle, the trip, the commodity quantity, the shipment, and the pickup and/or delivery operation.
personal mobility planning and urban engineering, those relationships are essential to deploy pertinent solutions and methods. Moreover, in personal mobility, where surveys are available, accessibility can be related directly to surveyed flows. However, in freight transport, there is a small number of surveys and a high difficulty of collecting suitable data (Holguín-Veras and Jaller, 2014; Gonzalez-Feliu, 2019). For those reasons, an analysis relating accessibility to estimates of FTG can be a valid alternative. A first work has been done in Gonzalez-Feliu and Peris-Pla (2017) relating FTG and STG (Shopping Trip Generation) of only retailing activities to a retailing attractiveness indication (i.e. a reciprocal measure to accessibility). That work raised also the issue of FTG data aggregation. Indeed, FTG is generally made at an establishment level, and few works relate FTG to the characteristics of urban zones. Indeed, to the best of our knowledge, Lawson et al. (2012) use land classifications for an FTG zonal model. Sánchez-Díaz et al. (2012) made a first attempt of comparing individual generation models with zonal models (those of Lawson et al., 2012), and Gonzalez-Feliu and Peris-Pla (2017) proposed a first generalization of FTG from individual to zonal generation but only for retail activities.

3. Methodology

The basic idea behind the proposed methodology is that it would be possible to estimate a relationship between FTG and accessibility in any city, using standard data and without having extended survey data to estimate freight trips. Such a methodology would be valuable since not all territories have the necessary funds to carry out extensive surveys (Holguín-Veras and Jaller, 2014). To accomplish this, the proposed methodology is structured in three phases. The first phase is that of FTG estimation. To do this, a FTG model based on the framework of Sánchez-Díaz and Gonzalez-Feliu (2019) is deployed. The second phase is that of gravity accessibility indicator estimation. In the proposed framework, two accessibility indexes are estimated: a potential and an exponential indicators. Finally, the third phase is the assessment of the relationships between FTG and accessibility using a linear regression assessment and the analysis of the results issued from this assessment.

The methodology can be summarized in the following steps:
1) Estimation of FTG rates: Individual FTG rates at the establishment level and aggregation of individual FTG rates at the zone level.
2) Estimation of accessibility (at zone level).
3) Assessment of the relations between FTG and accessibility.

3.1. Estimation of FTG rates

The estimation of FTG rates is a well-studied field in literature. Most authors state that the determinants and generation patterns in FTG are directly related to the nature and the intensity of each activity in a given zone (Watson, 1975; Ogden, 1992; Eriksson, 1997; D’Este, 2000; Holguín-Veras et al., 2011; Alho and de Abreu e Silva, 2014; Oliveira and Pereira, 2014; Sánchez-Díaz, 2016; Gonzalez-Feliu, 2019). In other words, to estimate the number of freight trips in a zone, it is necessary to identify the economic activities in that zone and generate those trips for each activity (Gonzalez-Feliu, 2019). The main difficulties in estimating FTG are in general those related to data availability (Holguín-Veras and Jaller, 2014). Since for a detailed estimation the monetary efforts needed are high, only a few cities have deployed detailed surveys, and they are not periodic, except commercial transport surveys in Germany that partially capture information on freight trips (Gonzalez-Feliu, 2019). For this reason, modelling can be a valid alternative. Moreover, since the aim of this exploratory paper is to examine the relationships between FTG and accessibility to make suitable estimations of FTG with standard data aggregated at zonal level, using a model can give pertinent estimations to feed those analyses. Since the analyses will be carried out in French cities and in that context various models and methods issued from the same databases are found, we propose to adapt an existing model to generate FTG trips. Indeed, several authors stated on the pertinence and transferability of FTG models based on French surveys (Aubert and Routhier, 1999; Ambrosini et al., 2008, 2013; Dablanc and Routhier, 2009; CERTU, 2013; Gardrat, 2013; Guerrero et al., 2014; Sánchez-Díaz and Gonzalez-Feliu, 2019). Moreover, the analyses in this paper will focus on the relations between FTG and accessibility at zonal levels and not on the generation patterns of FTG. Finally, the models proposed by Sánchez-Díaz and Gonzalez-Feliu...
on how the functional form of each category is obtained and on the suitability of aggregating and disaggregating data are found in Sánchez-Díaz and Gonzalez-Feliu (2019). We present here a model resulting in a different aggregation of data and results proposed by those authors, based on the most suitable individual FTG estimates. To choose the most suitable category aggregation, we have examined RMSE for each category, and eventually merged those that resulted on a lower RMSE when merged. At a zonal level, we can estimate the FTG rates by aggregation of individual rates, i.e. on trips generated by each establishment. In other words, given a zone $i$, the number of freight trips generated by the zone can be estimated as follows:

$$T_i = \sum_{k=1}^{m} \sum_{j \in V_i^k} n_{ij}$$

where $V_i^k$ is the set of establishments of category $k$ in zone $i$ and $n_{ij}$ the number of freight trips of establishment $j$ belonging to set $V_i^k$

<table>
<thead>
<tr>
<th>Macro-category</th>
<th>Category</th>
<th>Name</th>
<th>Constant</th>
<th>Employment</th>
<th>RMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.1</td>
<td>Agriculture</td>
<td>2.8</td>
<td>0.21</td>
<td>3.95</td>
</tr>
<tr>
<td>2</td>
<td>2.1</td>
<td>Craftsmen</td>
<td>3.19</td>
<td>1.01</td>
<td>9.66</td>
</tr>
<tr>
<td>2</td>
<td>2.2</td>
<td>Offices</td>
<td>3.57</td>
<td>0.02</td>
<td>7.8</td>
</tr>
<tr>
<td></td>
<td>2.3</td>
<td>Tertiary (non-offices) and services</td>
<td>4.63</td>
<td>-</td>
<td>14.68</td>
</tr>
<tr>
<td>3</td>
<td>3.1</td>
<td>Chemical industry</td>
<td>23.88</td>
<td>0.15</td>
<td>30.32</td>
</tr>
<tr>
<td>3</td>
<td>3.2</td>
<td>Construction industry</td>
<td>6.57</td>
<td>0.21</td>
<td>20.5</td>
</tr>
<tr>
<td></td>
<td>3.3</td>
<td>Primary and intermediate products</td>
<td>6.52</td>
<td>0.25</td>
<td>15.09</td>
</tr>
<tr>
<td>3</td>
<td>3.4</td>
<td>Food and non-fragile consumer goods</td>
<td>8.68</td>
<td>0.23</td>
<td>23.68</td>
</tr>
<tr>
<td>4</td>
<td>4.1</td>
<td>Wholesale</td>
<td>18.74</td>
<td>0.68</td>
<td>88.28</td>
</tr>
<tr>
<td>5</td>
<td>5.1</td>
<td>Department stores</td>
<td>-</td>
<td>0.54</td>
<td>44.47</td>
</tr>
<tr>
<td>6</td>
<td>6.1</td>
<td>Clothing, shoes, leather</td>
<td>2.01</td>
<td>0.17</td>
<td>2.19</td>
</tr>
<tr>
<td>6</td>
<td>6.2</td>
<td>Butcher's shops</td>
<td>3.55</td>
<td>1.18</td>
<td>5.85</td>
</tr>
<tr>
<td>6</td>
<td>6.3</td>
<td>Small groceries</td>
<td>4.34</td>
<td>1.02</td>
<td>8.03</td>
</tr>
<tr>
<td>6</td>
<td>6.4</td>
<td>Bakery retailers</td>
<td>7.31</td>
<td>-</td>
<td>7.37</td>
</tr>
<tr>
<td>6</td>
<td>6.5</td>
<td>Hotels, restaurants, cafés</td>
<td>2.63</td>
<td>0.61</td>
<td>7.14</td>
</tr>
<tr>
<td>6</td>
<td>6.6</td>
<td>Pharmacies</td>
<td>15.94</td>
<td>1.94</td>
<td>14.71</td>
</tr>
<tr>
<td>6</td>
<td>6.7</td>
<td>Hardware stores</td>
<td>2.1</td>
<td>0.87</td>
<td>3.69</td>
</tr>
<tr>
<td>6</td>
<td>6.8</td>
<td>Furnishing shops</td>
<td>6.11</td>
<td>0.1</td>
<td>8.51</td>
</tr>
<tr>
<td>6</td>
<td>6.9</td>
<td>Bookshops</td>
<td>10.25</td>
<td>-</td>
<td>7.03</td>
</tr>
<tr>
<td>6</td>
<td>6.10</td>
<td>Street trading (marketplaces)</td>
<td>5.77</td>
<td>-</td>
<td>21.01</td>
</tr>
<tr>
<td>6</td>
<td>6.11</td>
<td>Other retail shops</td>
<td>-</td>
<td>0.96</td>
<td>7.52</td>
</tr>
<tr>
<td>7</td>
<td>7.1</td>
<td>Only Transport</td>
<td>10.76</td>
<td>-</td>
<td>35.55</td>
</tr>
<tr>
<td>7</td>
<td>7.2</td>
<td>Transport and warehousing</td>
<td>-</td>
<td>0.04</td>
<td>33.2</td>
</tr>
</tbody>
</table>
3.2. Accessibility indicators

The gravity accessibility indicators taken into account here are derived from that of Hansen (1959). In personal transport, the opportunities are in general set up to the number of employees in the zone of destination (overall for work accessibility, retailing-based for shopping accessibility and tertiary/service-based for accessibility to services, etc.). Some authors define the opportunities as the quantity of freight generated by the zone of the destination (Crainic et al., 2010). However, the aim of this research is to relate a generated value (freight trips) and an indicator of accessibility (which is in general estimated with censorial data), so it seems more suitable to not deploy a Freight Generation (FG) model for the accessibility index estimation. Furthermore, FG and FTG rates have a relationship, even if it is not causal, so estimating FG on an indicator of accessibility will lead to non-independence between FTG and this index. Thus, taking this into account, we define the following potential accessibility index,

\[ A^P_i = \sum_{j=1}^{n} \text{Emp}_j \cdot d_{ij}^{-\alpha} \cdot 1000 \]  

where:
- \( \text{Emp}_j \) is the employment in the zone of the destination.
- \( d_{ij} \) is the Euclidean distance between those two zones. Indeed, in most works dealing with freight transport, transport cost is assimilated to distances.
- \( \alpha \) is a parameter set to 0.98, according to Gonzalez-Feliu’s (2019) considerations.

A variant of the potential accessibility is that of the exponential accessibility, defined for urban goods transport as follows:

\[ A^E_i = \sum_{j=1}^{n} \text{Emp}_j \cdot e^{-\beta d_{ij}} \cdot 1000 \]  

where:
- \( \text{Emp}_j \) is the employment in the zone of destination.
- \( d_{ij} \) is the Euclidean distance between those two zones.
- \( \beta \) is a parameter, set to 0.18, according to Gonzalez-Feliu’s (2019) considerations.

Parameters are set by defining those indicators based on shopping trips (which are urban goods transports of a different category) to get a pertinent indicator linking inbound and outbound flows related to retailers. In both cases, a scaling factor of 1000 is applied to make it easier for reading and understanding of accessibility indexes.

3.3. Assessment of the relations between FTG and accessibility

To assess the suitability of both accessibility indexes in predicting FTG, we propose first a descriptive statistics’ analysis of both accessibility indicators (in terms of dispersion and symmetry), then we estimate both relationships using linear regression. However, since other variables can be considered, we will relate the trip generation to a set of variables following a general trip generation rates relationship of type:

\[ T_i = a^0 + a^{dem} \cdot X_i^{dem} + a^{soc} \cdot X_i^{soc} + a^{acc} \cdot X_i^{acc} \]  

where \( X_i^{dem} \), \( X_i^{soc} \) and \( X_i^{acc} \) are the sets of demographic, socio-economic and accessibility variables respectively that characterize zone \( i \); \( a^{dem} \), \( a^{soc} \) and \( a^{acc} \) the sets of constant parameters associated to corresponding variables; and \( a^0 \) is a constant coefficient. To assess these relationships, we will carry out linear regression analyses, to test the independence of the considered variables but also to ensure that the main hypotheses behind linear regression are valid. Linear regression validity is subject to the verification of several hypotheses: mainly non-collinearity among explanatory variables, independence and exogeneity of error terms, homoscedasticity and normality of error terms (Wonnacott and Wonnacott, 2001). The normality of errors is not fundamental for ensuring the validity of the linear regression, but since the number of individuals (here the number of zones for the analysis of accessibility, see next section) used for the regression is higher than 35, we can assume it (Wonnacott and Wonnacott, 2001). Considering that those analyses remain standard and known, we will present them in detail in the following section, after defining the data used to carry out the analyses and the main variables considered.

4. Results

We present here the results of the linear regression analyses. We apply the proposed framework to the
French city of Lyon and its conurbation. The urban area considered in the analysis has about 2 million inhabitants, about 122 000 economic activities (i.e. establishments) and a little less than 830 000 employees. We present here an aggregation and the consequent analysis by zone. Indeed, the gravity accessibility being necessarily related to distance, a zonal aggregation is indispensable to produce pertinent indexes. Euclidean distances are estimated via interpolation of a French standard zoning file (IRIS) based on 83 zones (Gonzalez-Feliu, 2019). Finally, to each zone, a number of establishments and employees are associated by aggregating data of an establishment file of Lyon in the year 2005. A distance matrix is then estimated.

To do this analysis, we estimate first the FTG using the proposed model, then we estimate both the potential and the exponential accessibility as presented above.

We can moreover use other variables to explain the FTG rates $T_i$ (explained in Section 3.3). After observing the available data, we selected ten explanatory variables able to be considered for the analysis, related to the available data:

- $\text{POP}_i$: Population of zone $i$.
- $\text{EMP}_i$: Total employment for zone $i$ (for all activities).
- $\text{EST}_i$: Total number of establishments for zone $i$ (for all activities).
- $\text{SUR}_i$: Surface of zone $i$ (in km).
- $\text{DEP}_i$: Density of population for zone $i$.
- $\text{DES}_i$: Density of establishments for zone $i$.
- $\text{DEM}_i$: Density of employment for zone $i$.
- $\text{DIC}_i$: Distance of zone $i$ to the city centre.
- $\text{POT}_i$: Potential accessibility, estimated as presented above.
- $\text{EXP}_i$: Exponential accessibility, estimated as presented above.

Although other variables would be considered, we report here those that are able to be estimated from the available data presented above. Before providing a linear regression analysis, it is important to identify the potential of those variables to explain FTG. To do this we present, in Table 2 below, a collinearity test between FTG and each explanatory variable.

<table>
<thead>
<tr>
<th>Variable 1</th>
<th>Variable 2</th>
<th>Pearson coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td>FTG rates</td>
<td>Potential accessibility</td>
<td>0.73</td>
</tr>
<tr>
<td>FTG rates</td>
<td>Exponential accessibility</td>
<td>0.66</td>
</tr>
<tr>
<td>FTG rates</td>
<td>Surface</td>
<td>-0.11</td>
</tr>
<tr>
<td>FTG rates</td>
<td>Population</td>
<td>0.35</td>
</tr>
<tr>
<td>FTG rates</td>
<td>DistCen</td>
<td>-0.36</td>
</tr>
<tr>
<td>FTG rates</td>
<td>Employment</td>
<td>0.90</td>
</tr>
<tr>
<td>FTG rates</td>
<td>Number of establishments</td>
<td>0.82</td>
</tr>
<tr>
<td>FTG rates</td>
<td>Emp density</td>
<td>0.44</td>
</tr>
<tr>
<td>FTG rates</td>
<td>Est density</td>
<td>0.27</td>
</tr>
<tr>
<td>FTG rates</td>
<td>Pop density</td>
<td>0.18</td>
</tr>
</tbody>
</table>

We observe that linear relationships can be established with different variables. The highest collinearity between FTG rates and an exploratory variable is observed for the employment (0.90), followed by the number of establishments (0.82). Then, potential and exponential accessibilities present valid Pearson coefficients to consider linear relationships (0.73 and 0.66 respectively). All other coefficients have an absolute value lower than 0.5. Concerning the sense of the relationship, only two variables (surface and distance to the city centre) present negative values (which would result into inverse proportional relationships) but the values are too small to consider a potential linear regression with only one of those two variables. For all variables with a Pearson coefficient with an absolute value higher than 0.5, this coefficient is positive, which means that in those cases FTG are directly proportional to those variables.

After testing collinearity between FTG and each explanatory variable, it is important to also test collinearity between every two pairs of explanatory variables. However, since we aim to explore the links between FTG and accessibility mainly, we present in Table 3 only the correlation coefficients between each accessibility indicator and other explanatory variables.

According the Table 3, both accessibilities indexes are highly correlated (Pearson coefficient of 0.98) and we observe that they are also correlated to most other explanatory variables. Only the population and, in a lower order of magnitude, the surface and the population density seem to have a small linear correlation. However, those two last present Pearson
coefficients of about 0.5 for both accessibility indicators whereas those relating each accessibility indicator to population are close to zero, so we can consider accessibility and population as independent variables.

Table 3. Correlation analysis between each accessibility indicator and other explanatory variables

<table>
<thead>
<tr>
<th>Variable 1</th>
<th>Variable 2</th>
<th>Pearson coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td>Potential accessibility</td>
<td>Surface</td>
<td>-0.45</td>
</tr>
<tr>
<td></td>
<td>Population</td>
<td>0.04</td>
</tr>
<tr>
<td></td>
<td>DistCen</td>
<td>-0.56</td>
</tr>
<tr>
<td></td>
<td>Employment</td>
<td>0.80</td>
</tr>
<tr>
<td></td>
<td>Number of establishments</td>
<td>0.65</td>
</tr>
<tr>
<td></td>
<td>Emp density</td>
<td>0.83</td>
</tr>
<tr>
<td></td>
<td>Est density</td>
<td>0.64</td>
</tr>
<tr>
<td></td>
<td>Pop density</td>
<td>0.50</td>
</tr>
<tr>
<td>Exponential accessibility</td>
<td>Surface</td>
<td>-0.51</td>
</tr>
<tr>
<td></td>
<td>Population</td>
<td>-0.01</td>
</tr>
<tr>
<td></td>
<td>DistCen</td>
<td>-0.61</td>
</tr>
<tr>
<td></td>
<td>Employment</td>
<td>-0.69</td>
</tr>
<tr>
<td></td>
<td>Number of establishments</td>
<td>0.58</td>
</tr>
<tr>
<td></td>
<td>Emp density</td>
<td>0.85</td>
</tr>
<tr>
<td></td>
<td>Est density</td>
<td>0.68</td>
</tr>
<tr>
<td></td>
<td>Pop density</td>
<td>0.57</td>
</tr>
<tr>
<td>Potential accessibility, Exponential accessibility</td>
<td>-0.98</td>
<td></td>
</tr>
</tbody>
</table>

Therefore, we will only consider as regression variables accessibility (potential and exponential) and population. Concerning the use of employment (EMP) as an explanatory variable, the FTG modelling framework uses, in many categories, that variable (at the individual level) so the correlation with accessibility is important. Since we aim to analyse the links between accessibility and FTG and employment has a non-negligible correlation to accessibility, that variable will not be considered in the following assessments. We can then define the relationship between FTG and the considered explanatory variables as follows:

\[
FTG_i = A_0 + A_1 \cdot POT_i + A_2 \cdot EXP_i + A_3 \cdot POP_i
\]

Knowing that \(POT_i\) and \(EXP_i\) have a collinearity, the models containing both accessibilities at the same time are not examined.

We present in Table 4 below the results of the regression analysis. We report in the first column the name of the explanatory variables used in each regression, and in columns two to five the values of the concerned coefficients. In the remaining 4 columns, we report respectively the R² value, the F-Test result, the T-Test result and the result of a Durbin-Watson error correlation test (for the four assessments involving two or more variables, since error correlation between explanatory variables is only present when two or more of those variables are defined).

Out of eight regressions tested, six passed the three proposed tests. We observe that Durbin-Watson test is verified in all four assessments involving both an accessibility measure and population as explanatory variables but the T-Test is negative in those considering the constant term. For relationships including only an accessibility measure as an explanatory variable, all considered tests are positive (and no correlation test is necessary since only one variable is defined). If we compare those relationships passing all tests, we observe that those with only one variable have an R² lower than those with both an accessibility variable and the population, and the difference is significant (more than 0.1 between the best relationship with one explanatory variable and the less performing of those with two explanatory variables).

Table 4. Regression indicators for the different possibilities

<table>
<thead>
<tr>
<th>Variables</th>
<th>(A_0)</th>
<th>(A_1)</th>
<th>(A_2)</th>
<th>(A_3)</th>
<th>(A_4)</th>
<th>R²</th>
<th>F-Test</th>
<th>T-Test</th>
<th>Durbin-Watson test</th>
</tr>
</thead>
<tbody>
<tr>
<td>POT</td>
<td>6658.06</td>
<td>67.53</td>
<td></td>
<td></td>
<td></td>
<td>0.53</td>
<td>OK</td>
<td>OK</td>
<td>Not applicable</td>
</tr>
<tr>
<td>POT</td>
<td>115.52</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.77</td>
<td>OK</td>
<td>OK</td>
<td>Not applicable</td>
</tr>
<tr>
<td>EXP</td>
<td>7940.07</td>
<td>45.19</td>
<td></td>
<td></td>
<td></td>
<td>0.43</td>
<td>OK</td>
<td>OK</td>
<td>Not applicable</td>
</tr>
<tr>
<td>EXP</td>
<td></td>
<td>88.02</td>
<td></td>
<td></td>
<td></td>
<td>0.66</td>
<td>OK</td>
<td>OK</td>
<td>Not applicable</td>
</tr>
<tr>
<td>POT, POP</td>
<td>2613.88</td>
<td>66.07</td>
<td>0.20</td>
<td></td>
<td></td>
<td>0.63</td>
<td>OK</td>
<td>Not for constant Independent</td>
<td></td>
</tr>
<tr>
<td>POT, POP</td>
<td></td>
<td>72.05</td>
<td>0.28</td>
<td></td>
<td></td>
<td>0.90</td>
<td>OK</td>
<td>OK</td>
<td>Independent</td>
</tr>
<tr>
<td>EXP, POP</td>
<td>3249.23</td>
<td>45.50</td>
<td>0.22</td>
<td></td>
<td></td>
<td>0.55</td>
<td>OK</td>
<td>Not for constant Independent</td>
<td></td>
</tr>
<tr>
<td>EXP, POP</td>
<td></td>
<td>50.78</td>
<td>0.33</td>
<td></td>
<td></td>
<td>0.88</td>
<td>OK</td>
<td>OK</td>
<td>Independent</td>
</tr>
</tbody>
</table>
Moreover, for those two last relationships, the R² is close to 0.9 in both cases, which remains very high. Thus, adding the population as an explanatory variable in addition to accessibility increases the quality of the regression model. Moreover, both regressions verified the last hypothesis to verify, that of Homoscedasticity (Wonnacott and Wonnacott, 2001). Since the R² is close for both regressions (i.e. the one with POT and POP as explanatory variables, R²=0.90 and the one with EXP and POP, R²=0.88), we chose to retain the two relationships. Thus both of them can be considered as valid:

\[
T_i = 72.05 \cdot POT_i + 0.28 \cdot POP_i \\
\text{or} \\
T_i = 50.78 \cdot EXP_i + 0.33 \cdot POP_i
\]

Those two relationships can be also used to define estimations of freight trips from standard information. Indeed, population is a known information in many cities, and international databases allow estimating it in a quite satisfactory way, at a more than enough disaggregation scales (1km² areas covering most territories, Winkenbach et al., 2018). Moreover, the definition of both the potential and the exponential accessibility indicators use standard data: employments as opportunities, which can be retrieved for main cities at National Statistics or Chamber of Commerce registries (Gonzalez-Feliu, 2019). Furthermore, those models remain coherent, in terms of structure and data needs, to those of personal mobility, and then can be used, with the same accuracy, to have a quick estimate of freight flows to include into global freight and passenger transport models for traffic and civil engineering planning needs (Arvidsson et al., 2016; Pimentel and Alvelos, 2018).

These results have several research and practical implications. The first is that this correlation leads to the definition of freight trip model that does not need specific surveys to be deployed, and can then be deployed with small costs, in an analogous way that what is done in passenger transport (Crozet et al., 2012). Moreover, this framework can be applied and understood easily, without a need for specialized engineers and costly assistance. This leads to a first group of research works, related to the evolution of current freight transport models (which have in most cases an establishment-based generation, even if some of them make an aggregation of results to have zonal O/D matrices) to produce more coherent data in an aim of comparing passenger and freight transport models. This would lead into the use, in practice, of joint trip generation rates of zones for traffic assessment (mainly at the level of streets or crossroads) as well as to feed traffic micro-simulation approaches in input data (after converting the generated average FTG into random-based probabilistic estimations, extending the work of Chiabaut et al., 2016).

A second implication derives from the fact that the accessibility gravity measurement can also be a first approximation to distribute trips and produce a raw estimation of O-D matrices. However, this approximation will not be as accurate as an ad-hoc estimation using more performant methods, but can be applied with small amounts of data and does not need count records for model calibration. Finally, and if we consider the transferability hypothesis of FTG rates (Holguin-Veras et al., 2013), it can be applied in any context, if a good calibration of a transferable FTG model is made and a consequent deployment of the linear model follows. In that context, simple freight trip models, with levels of accuracy that can be accepted for overall estimations of freight flows at macroscopic scales like a city or a conurbation, can be developed and used for both urban planning or logistics strategic management. Those models can, after completing FTG rates by a quantity of freight per delivery, to feed vehicle routing approaches and propose complementary data input sets to operations research communities. They can also be combined to operational tool (Erdogan, 2017) to make estimates of routes, being indicated in practice for public and private policymaking in terms of logistics distribution (access restrictions, parking planning, zonal definitions of delivery services, etc.)

Main practical uses of the accessibility indexes are related to the economic development of zones (identification of potential zones of development and their influence to truck traffic), linking economic activities and freight needs (via conversion factors to estimate freight quantities) or identifying inequalities in terms of goods transport and logistics, among others. Indeed, the main implication of the results shown above is that FTG rates can be estimated at a zonal level knowing the population (which is a well-known information in most cities) and an accessibility indicator (potential or exponential) estimated using the total number of employees per zone and the
Euclidean distances between each two zones (both of them being easy to estimate with data available in most cities). Estimating FTG rates at the individual establishment level needs instead detailed individual information (Holguin-Veras et al., 2013) and according to several authors, it is necessary to carry out detailed (and expensive) data collection campaigns (Ambrosini et al., 2008; Holguin-Veras and Jaller, 2014).

The results presented here show that using accessibility and population (which are easy to estimate with no need of collecting new data), the results obtained (at zonal levels), are very close to those of individual FTG models aggregated at a zonal level (R² of 0.88 and 0.9 respectively for exponential and potential accessibility). So urban planners can, with this framework, estimate FTG rates with data currently available for them with a suitable accuracy. Moreover, since accessibility is a well-known notion for urban planners, the proposed framework allows to include freight flows in urban plans or transport engineering studies with little efforts for planners, who will not need to deploy specific complex models and tools and use existing ones.

However, those results apply nowadays only to France, for which the transferability of FTG models is verified (Gonzalez-Feliu, 2019). US models were also verified (Holguin-Veras et al., 2013), and those analyses would be replicated with US data to deploy an analogous framework for the US context. In any case, it seems that if basic FTG models are available in a country and the transferability of FTG models from one city to another is verified, the proposed framework here is then able to be deployed and used.

5. Conclusion

This paper is a first step in exploring the relationships between FTG and freight accessibility. It is also one of the first works that compare two freight accessibility indicators and their possibilities into linking them to freight demand estimation (in terms of trip generation). The potential accessibility seems to be more closely related to FTG and the resulting relationships seem to be a good way to approximate FTG when few or no data of specific trip surveys are available. However, this work makes a first approximation, i.e. the FTG are estimated and not surveyed, which is valid when it is not possible to deploy such surveys (due to restrictions on costs or other resources). However, it shows that FTG rates can be related to the land accessibility as well as to the population. The main practical implications of the work are related to the deployment of models when few data is available. The proposed framework allows us to estimate FTG rates in a quite accurate way (R² close to 0.9) with only two sources of information. The first, population, is more or less able to be estimated in most urban zones. Moreover, Population can be obtained, for a small granularity (grids of 1 km²) at the LandScan™ database, for any city of the World (Bright et al., 2018 and Regal, 2020). The second, the accessibility indicator, is calculated using employment (able to be estimated in a more or less accurate way in many cities) and distances (for which many estimations can be made). Thus, the proposed framework uses standard data and it is easy to assess with current information, making it a first approach to overcome the main difficulties of getting very detailed data without losing accuracy, which is very important. This work remains then exploratory but gives a first idea of the potential of using accessibility related to trip generation in urban logistics. The main implication of the proposed work is practical: using standard available data, urban planners can estimate FTG rates with a good accuracy and no need of deploying specific and expensive data collection campaigns/surveys. Those relationships seem relevant for two reasons. First, those results allow defining simplified zonal models that can be useful for quick estimates of FTG rates at urban level using little amounts of standard data. Second, the estimation of FTG at a zonal level can include variables not able to be considered at the establishment level, like population on various density measures, among others, and be more coherent with personal trip and accessibility modelling approaches, supporting the integrations of both types of trips. Nevertheless, and since population and accessibility (the two most significant explanatory variables) are able to be estimated with available data in most cities, the proposed framework remains a potential tool able to be used and understood by city engineers and planners in practice.

The proposed framework is preliminary and can be improved. More-in-depth validations seem however necessary to generalize this work into a standard framework. For that reason, further works will focus on two main elements: the first is that of estimating the quality of the FTG model and that of the proposed accessibility-based generation procedures.
concerning surveyed data; the second concerns the indicator itself, which has not been analyzed here based on accessibility representation. Indeed, those indicators are traditionally defined on the basis of personal transport, with the hypothesis that households have a strong influence on daily trips. But because the nature of freight trip chains and the spatial distribution of the different generators are essentially different than that of the personal trip generation determinants, a freight-specific indicator would have a better relationship to FTG. In this sense, future work will examine more in-depth those determinants in order to define more suitable indicators for freight and logistics accessibility.

Acknowledgements
This research was initially done during Jesus González-Feliu’s and Laura Palacios-Argüello’s positions at Ecole des Mines de Saint-Etienne. Part of this work has been carried out in the context of France-South American project DEMOSDUM from the Math AmSud 2015 program.

References


[37] HOLGUÍN-VERAS, J., RAMÍREZ-RÍOS, D.G., GONZÁLEZ-FELIU, J., CASPERSEN,


[56] VAN DEN HEUVEL, F. P., RIVERA, L., VAN DONSELAAR, K. H., DE JONG, A., SHEFFI,


